AMENDMENTS TO THE CLAIMS:

This listing of claims will replace all prior versions, and listings, of claims in the application:

LISTING OF CLAIMS:

1-21. (canceled)

22. (currently amended) A method for gene mapping to locate a gene associated with a certain phenotype from a dataset of chromosome and phenotype data ~~from a database, comprising analyzing linkage disequilibrium between~~ by analyzing an association between phenotype and genetic ~~marks~~ markers m_i, comprising:

i) searching from ~~the data~~ said dataset for all marker patterns P that satisfy a pattern evaluation function $e(P)$, wherein

a: the marker patterns are expressions within ~~the database~~ said dataset comprising genetic markers and their alleles and zero or more of the following: individual covariates, environmental variables and auxiliary phenotypes; and

b: the pattern evaluation function $e(P)$ is a measure of the association between the marker pattern P and a phenotype being studied,

ii) scoring each marker m_i of the data with a marker score $s(m_i)$, which is a function of the set S_i defined as the set of marker

patterns overlapping the marker m_i and satisfying the pattern evaluation function e as defined in step i), and

iii) locating said gene to the marker m_i having the best score $s(m_i)$, wherein the best score is the highest obtained score if said mapping the location of a gene by evaluating the scores $s(m_i)$ of all the markers m_i in the data which is determined by maximizing the score if [[the]] said scoring function is designed to give higher scores closer to the gene, ~~and on minimizing the score if the scoring function is designed to give lower scores closer to the gene~~ on locating said gene to a chromosomal region containing a set of best scoring markers.

23. (previously presented) The method of claim 22, wherein the chromosome data consists of either haplotypes or genotypes.

24. (previously presented) The method of claim 23, wherein said haplotypes and genotypes contain flexible regions.

25-29. (canceled)

30. (currently amended) The method of claim 22, wherein

a) the phenotype being studied is qualitative, and

b) the pattern evaluation function $e(P)$ has the form $e(P)$ =

true if and only if $e'(P) > x$, where $e'(P)$ is the ~~(signed)~~ signed association measure χ^2 and x is a user specified minimum value wherein said signed value of the χ^2 is negative if the relative frequency of the halotype pattern among the control chromosomes is higher than that of the trait-associatead chromosomes, and otherwise positive, and

c) the score $s(m_i)$ of marker m_i is the size of S_i, also called marker-wise pattern frequency of m_i and denoted by $f(m_i)$.

31. (previously presented) The method of claim 22, wherein

a) the pattern evaluation function $e(P)$ has the form $e(P)$ = *true if and only if* $e'(P) > x$, where $e'(P)$ is the absolute frequency of pattern P in the data and x is a user-specified value,

b) in order to derive the score $s(m_i)$, the p value (statistical significance) of each marker pattern P in determining the phenotype being studied is evaluated, and

c) the score $s(m_i)$ is the distance between the observed p value distribution of patterns in S_i and the uniform distribution, defined as average of $(p_i - q_i) \log (p_i / q_i)$ over all $i = 1..n$, where n is the number of haplotype

patterns in S_i, p_i is the ith smallest p value in S_i, and q_i is the expectation of the ith smallest p value, if the p values were randomly drawn from the uniform distribution.

32. (previously presented) The method of claim 31, where the p value is computed using a linear model of form $Y = \beta_1 X_1 + \ldots + \beta_k X_k + \alpha Z + \beta_0$, where the dependent variable Y is the phenotype being studied, X_1 through X_k are covariates, and Z is a dummy variable for the occurrence of the haplotype pattern, and the coefficients α and β_* are adjusted for best fit, and then the significance of Z as a covariate is assessed using a t test with the null hypothesis "$\alpha = 0$".

33. (previously presented) The method of claim 22, further refining each score $s(m_i)$ by replacing it by the marker-wise p value of the score $s(m_i)$, where the statistical significance of $s(m_i)$ is measured against the null hypotheses that there is no gene effect.

34. (previously presented) The method of claim 22, wherein an area returned from a prediction of a gene location is contiguous or fragmented or a point.

35-36. (canceled)

37. (previously presented) The method of claim 22, wherein the location of the gene, predicted as a function of the scores $s(m_i)$ and based on maximizing or minimizing the score, is determined by evaluating the marker scores or by visualization.

38. (canceled)

39. (previously presented) A computer-readable data storage medium having computer-executable program code stored thereon operative to perform the method of claim 22 when executed on a computer.

40. (previously presented) A computer system having executable program code that performs the method of claim 22.

41. (new) The method of claim 22, comprising searching patterns P by the following algorithm:

Input:

- set U of marker patterns
- evaluation function $e(P)$ for patterns P in U
- generalization relation $<$ for patterns in U, where function e and relation $<$ are such that if $e(P)$ is true and $P' < P$, then $e(P')$ is also true

Output:

- set S of patterns P in U satisfying $e(P)$

Definition:

- function Lss: $U \rightarrow 2^U$, $Lss(P) = \{ P' \in U \mid P < P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P < P'' < P' \}$ is the set of *least special specializations* of pattern P.

Method:

Initialize set S and set E of evaluated patterns to be empty sets.

Let *Gen* be the set of patterns P in U, for which there are no patterns P' in U : $P' < P$.

Recursively evaluate each pattern P in *Gen* in depth-first order.

Method for recursively evaluating pattern P :

Insert P into set E.

If $e(P)$ is true, then {

Insert P into set S.

Find set $Spec = \{ P' \in Lss(P) \mid P' \notin E \}$.

Recursively evaluate each pattern in *Spec*.

}

42. (new) The method of claim 22, comprising searching patterns P by the following algorithm:

Input:

- set U of marker patterns
- evaluation function $e(P)$ for patterns P in U
- frequency threshold x
- generalization relation $<$ for patterns in U, where relation $<$ is such that if $P' < P$, then occurrence of pattern P implies occurrence of pattern P'.

Output:

- set $S = \{P \in U \mid e(P)$ and $ae(P)$ are true$\}$ of patterns, where $ae(P)$ is true if and only if the frequency of pattern P exceeds a given threshold x

Definition:

- function Lss: $U \rightarrow 2^U$, $Lss(P) = \{ P' \in U \mid P < P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P < P'' < P'\}$ is the set of *least special specializations* of pattern P.

Method:

Method:

Initialize set S and set E of evaluated patterns to be empty sets.

Let *Gen* be the set of patterns P in U, for which there are no patterns P' in U : $P' < P$.

Recursively evaluate each pattern P in *Gen* in depth-first order.

Method for recursively evaluating pattern P :

Insert P into set E.

if $ae(P)$ is true, then {

if $e(P)$ is true then {

insert P into set S.

}

Find set $Spec = \{ P' \in Lss(P) \mid P' \notin E \}$.

Recursively evaluate each pattern in *Spec*.

}

43. (new) The method of claim 22, comprising searching patterns P by the following algorithm:

Input:

- marker map $M = (m_1, \ldots, m_k)$
- phenotype vector $Y = (y_1, \ldots, y_n)$, where $y_i \in$ {'disease-associated','control'} for all i
- $n \times k$ haplotype matrix $H = \{h_{ji}\}$, where h_{ji} is the allele at

marker i in haplotype j

- threshold value x for χ^2 test for association between phenotype and pattern
- maximum pattern length l
- maximum number of gaps g
- maximum gap size s

Output:

- set S of patterns P in U satisfying $e(P)$, where U consists of marker-allele assignments that adhere to parameters l, g, and s, and $e(P)$ is true if and only if χ^2 test on P using haplotype matrix H and phenotypes Y exceeds given threshold x

Definitions:

- pattern $(p_1, \ldots, p_k)$ *matches* haplotype j, if and only if for all markers i: $p_i = h_{ji}$ or $p_i = *$
- *frequency* of pattern P is the number of haplotypes in H matched by P
- a *gap* in pattern $(p_1, \ldots, p_k)$ is set $\{i, \ldots, j\}$ of markers, $i \leq j$ and $i > 1$ and $j < k$, for which $p_{i-1} \neq *$ and $p_{j+1} \neq *$ and $p_n = *$ for each $n \in \{i, \ldots, j\}$.

Method:

Initialize set S to be empty set.

Calculate lower bound *lb* for pattern frequency: $lb = \pi_A\pi\ x\ /\ (\pi_C\pi + \pi_A x)$, where π_A is

the number of disease-associated haplotypes, π_C is the number of control haplotypes,

and π is $\pi_A + \pi_C$.

Initialize $(p_1, \ldots, p_k)$ to be empty pattern $(*, \ldots, *)$.

For each $i \in \{1, \ldots, k\}$ and for each $a \in A_i$, where A_i is the set of alleles at marker i: {

Let $p_i := a$.

Recursively evaluate pattern $(p_1, \ldots, p_k)$ ranging from i to i and all its extensions

to the right.

Let $p_i := *$.

}

Method for recursively evaluating pattern $(p_1, \ldots, p_k)$ ranging from i to j :

If χ^2 statistic for pattern $(p_1, \ldots, p_k)$ is greater than or equal to x and $p_j \neq *$, then

insert P into set S.

If $j < k$ and $j - i + 1 < l$ and the frequency of pattern $(p_1, \ldots, p_k)$ is greater than or

equal to *lb*, then {

For each possible allele a at marker $j+1$, $a \in A_{j+1}$: {

Let p_{j+1} : = a.

Recursively evaluate pattern (p_1, ..., p_k) ranging from i to j +1.

}

Let p_{j+1} : = *.

If $p_i \neq$ * and the number of gaps in pattern (p_1, ..., p_k) is smaller than g, then

introduce a new gap starting at marker j+1, and recursively

evaluate pattern (p_1, ..., p_k) ranging from i to j+1.

If p_i = * and the number of adjacent markers $j' \leq j$: $p_{j'}$ = * is smaller than s, then

extend the current gap over marker j+1, and recursively evaluate

pattern (p_1, ..., p_k) ranging from i to j+1.

}

44. (new) The method of claim 22, comprising searching patterns P by the following algorithm:

Input:

- set U of marker patterns
- evaluation function $e(P)$ for patterns P in U
- generalization relation < for patterns in U, where function e

and relation $<$ are such that if $e(P)$ is true and $P' < P$, then $e(P')$ is also true

Output:

- set S of patterns P in U satisfying $e(P)$

Definitions:

- function Lgg: $U \rightarrow 2^U$, $Lgg(P) = \{ P' \in U \mid P > P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P > P'' > P'\}$ is the set of *least general generalizations* of pattern P.
- function Lss: $U \rightarrow 2^U$, $Lss(P) = \{ P' \in U \mid P < P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P < P'' < P'\}$ is the set of *least special specializations* of pattern P.

Method:

Initialize set S and set Q to be empty sets.

Initialize set F to contain patterns P in U, for which there are no patterns P' in U : $P' < P$.

Repeat the following steps while F is a non-empty set :

For each $P \in F$: {

if $e(P)$ = true, then insert P into set S,

else remove P from set F.

}

Let $Q := Q \cup F$.

Initialize set C to be empty set.

For each $P \in F$:

Let $C := C \cup \{P' \in U \mid P' \in Lss(P)$ and for all $P'' \in Lgg(P') : P'' \in Q\}$.

Let $F := C$.

45. (new) The method of claim 22, wherein marker patterns P are searched by the following algorithm:

Input:

- set U of marker patterns
- evaluation function $e(P)$ for patterns P in U
- frequency threshold x
- generalization relation $<$ for patterns in U, where relation $<$ is such that if $P' < P$, then occurrence of pattern P implies occurrence of pattern P'.

Output:

- set $S = \{P \in U \mid e(P)$ and $ae(P)$ are true$\}$ of patterns, where $ae(P)$ is true if and only if the frequency of pattern P exceeds a given threshold x

Definitions:

- function *Lgg*: $U \rightarrow 2^U$, $Lgg(P) = \{ P' \in U \mid P > P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P > P'' > P'\}$ is the set of *least general generalizations* of pattern P.
- function *Lss*: $U \rightarrow 2^U$, $Lss(P) = \{ P' \in U \mid P < P'$ and $P' \neq P$ and there is no P'' such that $P'' \neq P$ and $P'' \neq P'$ and $P < P'' < P'\}$ is the set of *least special specializations* of pattern P.

Method:

Initialize set S and set Q to be empty sets.

Initialize set F to contain patterns P in U, for which there are no patterns P' in U : $P' < P$.

Repeat the following steps while F is a non-empty set :

For each $P \in F$: {

if $ae(P)$ = true, then {

if $e(P)$ is true, then insert P into set S.

}

else remove P from set F.

}

Let Q : = $Q \cup F$.

Initialize set C to be empty set.

For each $P \in F$:

Let $C := C \cup \{P' \in U \mid P' \in Lss(P)$ and for all $P'' \in Lgg(P'): P'' \in Q\}$.

Let $F := C$.

46. (new) The method of claim 22, wherein the location of the gene, predicted as a function of the scores $s(m_i)$ and based on maximizing or minimizing the score, is predicted by the combination of most probable intervals for containing a trait-susceptibility locus that covers a proportion t ranging from 0 to 100 % of the region covered by markers m_i obtained by taking all such points in said region whose nearest marker is within the k best scoring markers, and wherein k is selected so that the resulting area has length at most t times the length of said region.

47. (new) The method of claim 22, wherein the location of the gene, predicted as a function of the scores $s(m_i)$ and based on maximizing or minimizing the score, is predicted to those points in the region covered by markers m_i whose nearest marker scores …

48. (new) The method of claim 22, comprising searching for multiple genes using patterns comprising several components referring to different potential gene loci, and scoring marker tuples instead of plain markers.